Predicting failure of hematopoietic stem cell mobilization before it starts: the Predicted Poor Mobilizer (pPM) score

TITLE:

Predicting failure of hematopoietic stem cell mobilization before it starts: the Predicted Poor Mobilizer (pPM) score.

RUNNING HEAD: PREDICTING FAILURE OF HEMATOPOIETIC STEM CELL MOBILIZATION

AUTHORS:

Jacopo Olivieri,[1,19] Immacolata Attolico,[2] Roberta Nuccorini,[2] Sara Pasquina Pascale,[2] Martina Chiarucci,[1] Monica Poiani,[1] Paolo Corradini,[3] Lucia Farina,[3] Gianluca Gaidano,[4] Luca Nassi,[4] Simona Sica,[5] Nicola Piccirillo,[5] Pietro Enrico Pioltelli,[6] Massimo Martino,[7] Tiziana Moscato,[7] Massimo Pini,[8] Francesco Zallio,[8] Fabio Ciceri,[9] Sarah Marktel,[9] Andrea Mengarelli,[10] Pellegrino Musto,[11] Saveria Capria,[12] Francesco Merli,[13] Katia Codeluppi,[13] Giuseppe Mele,[14] Francesco Lanza,[15] Giorgina Specchia,[16] Domenico Pastore,[16] Giuseppe Milone,[17] Francesco Saraceni,[15] Elvira Di Nardo,[18] Paolo Perseghin,[6] and Attilio Olivieri.[1]

AFFILIATIONS:

1. Clinica di Ematologia, Università Politecnica delle Marche, Ancona, Italy;
2. Ematologia-Azienda Ospedaliera San Carlo, Potenza, Italy;
3. Dipartimento di Ematologia e Oncoematologia pediatrica, Fondazione IRCCS Istituto Nazionale dei Tumori, Milano, Italy;
4. Department of Translational Medicine , University of Eastern Piedmont , Novara , Italy;
5. UOC Ematologia, Università Cattolica del Sacro Cuore, Policlinico Agostino Gemelli, Roma, Italy;
6. Clinica Ematologica, A.O. San Gerardo, Monza, Italy;
7. Hematology and Stem Cell Transplant, Azienda Ospedaliera BMM, Reggio Calabria, Italy;
8. Ematologia, AON SS Antonio e Biagio e Cesare Arrigo, Alessandria, Italy;
9. Ospedale San Raffaele, Haematology and BMT, Milano, Italy;
10. Ematologia –IFO Istituto Nazionale Tumori Regina Elena, Roma, Italy;
11. IRCCS, Centro di Riferimento Oncologico della Basilicata, Rionero in Vulture, Italy;
12. Ematologia, Università sapienza, Roma, Italy;
13. Arcispedale Santa Maria Nuova – IRCCS, Reggio Emilia, Italy;
14. UOC di Ematologia e Unità Trapianti, Osp. Antonio Perrino, Brindisi, Italy;
15. Hematology and Stem Cell Transplant, Ravenna Hospital, Ravenna, Italy;
16. UO Ematologia con Trapianto, AOU Policlinico Consorziale, Bari, Italy;
17. Unità Trapianto di Midollo Osseo, Dipartimento di Ematologia, Azienda Ospedaliera Policlinico Vittorio Emanuele, Catania, Italy;
18. Dipartimento di matematica "G. Peano" Università di Torino, Italy;
19. UOC Medicina Interna ed Ematologia, ASUR – AV3, Civitanova Marche, Italy.

CORRESPONDING AUTHOR:

Prof. Attilio Olivieri
E-mail: a.olivieri@univpm.it
Dept Haematology-Ospedali Riuniti di Ancona
Università Politecnica delle Marche-DiscliMo
Via Conca,71 Ancona-Italy
Tel +390715964226
Fax +390715964222

MANUSCRIPT METRICS: Abstract 200 words; Manuscript 4014 words. Tables 7; Figures 1; References 34
Supplementary file 1.

ABSTRACT

Predicting mobilization failure before it starts may enable patient-tailored strategies. Although consensus criteria for predicted PM (pPM) are available, their predictive performance has never been measured on real data. We retrospectively collected and analyzed 1318 mobilization procedures performed for MM and lymphoma patients in the plerixafor era. In our sample, 180/1318 (13.7%) were PM. The score resulting from published pPM criteria had sufficient performance for predicting PM, as measured by AUC (0.67, 95%CI: 0.63-0.72). We developed a new prediction model from multivariate analysis whose score (pPM-score) resulted in better AUC (0.80, 95%CI: 0.76-0.84, p<0001). pPM score included as risk factors: increasing age, diagnosis of NHL, positive bone marrow biopsy or cytopenias before mobilization, previous mobilization failure, priming strategy with G-CSF alone or without upfront plerixafor. A simplified version of pPM-score was categorized using a cut-off to maximize positive likelihood ratio (15.7, 95%CI: 9.9–24.8); specificity was 98% (95%CI: 97%-98.7%), sensitivity 31.7% (95%CI: 24.9%-39%); positive predictive value in our sample was 71.3% (95%CI: 60%-80.8%). Simplified pPM-score can “rule in” patients at very high risk for PM before starting mobilization, allowing changes in clinical management, such as choice of alternative priming strategies, to avoid highly likely mobilization failure.

Introduction

High dose chemotherapy followed by autologous stem cell rescue is a mainstay of treatment for Multiple Myeloma (MM), Non Hodgkin Lymphoma (NHL) and Hodgkin Disease (HD). Autologous stem cell transplant (auto-SCT) is almost exclusively performed today with peripheral blood stem cells (PBSCs) infusion; [1] therefore stem cell mobilization (SCM) currently represents a crucial step of the whole transplant process. A threshold of $2x10^6$CD34+/Kg is regarded by most centers as the minimum amount of PBSCs to be infused to in order to safely perform the auto-SCT procedure.[2]

Despite developments in SCM protocols, a proportion of patients between 5% and 30% fail to collect an adequate number of CD34+.[3,4,5,6,7] Poor mobilization forces the patient to undergo a re-mobilization procedure and in some cases leads to postponing or even abandoning a transplant strategy. Several factors have been associated with poor mobilization,[4,5,8,9,10,11,18,12] however a thorough profile of the patient at high risk of sub-optimal SCM is still missing.

The Gruppo Italiano Trapianto Midollo Osseo (GITMO) has recently proposed a definition of the '*proven poor mobilizer* (PPM) and the *predicted poor mobilizer* (pPM), adopting a consensus based on an analytic hierarchy process (AHP).[13] While the GITMO definition of the PPM appears straightforward and is currently adopted by most European centers, identification of pPM appears more nuanced, and the GITMO pPM criteria should be validated in clinical trials and common clinical practice.

Early identification of mobilization failure is even more important nowadays, given the availability of interventions to boost or rescue low-performing procedures, such as the CXCR4 antagonist plerixafor. Currently, low circulating CD34+ count before apheresis is widely accepted as the stronger parameter able to predict mobilization failure. Thus, to assist the clinician in a timely and cost-effective use of Plerixafor, various algorithms were developed, based on the circulating

CD34+ at day 4 (in case of steady-state mobilization) or at the time of white blood cell (WBC) recovery (in case of chemo-mobilization).[14,15] However, such algorithms are applicable belatedly, only a few hours before the apheresis procedure begins. Ideally, identification of patients at high risk of inadequate SCM should be performed before starting the mobilization process, and protocol planning should be individualized according to patient and disease characteristics, and to stem cell target dose. Such tailored approach might help to optimize resources management, avoiding suboptimal stem cell collection, need for re-mobilization, and redundant days of apheresis.

We therefore conducted this retrospective study with the aim to validate the predictive ability of GITMO criteria for pPM, by measuring their diagnostic accuracy for the outcome of mobilization failure. Furthermore, by analyzing SCM kinetics in a large cohort of myeloma and lymphoma patients, we aimed to improve their predictive ability by adding new data, in order to elaborate a "poor mobilization risk score" easily applicable in the everyday practice, to help decision-making and procedure customization based on pre-mobilization parameters.

Methods

This was a multicenter retrospective observational study involving 17 Italian GITMO centers. The protocol was approved by the Ethics Committee of Potenza and subsequently by all participating centers. A waiver of patient's informed consent was obtained, owing to the retrospective nature of the study and provided that all patients' data were collected and managed after being anonymized. The study was conducted in accordance to Helsinki declaration, Good Clinical Practice and of applicable national regulations. All Centers were asked to fill a database containing informations on all mobilization attempts performed between Jan 1st 2009 and Jan 31st 2014 in patients with Multiple Myeloma (MM), Hodgkin's (HL) and non-Hodgkin's Lymphomas (NHL).

Collected data pertained to patient's characteristics, underlying hematological disease, therapeutic history before mobilization and kinetics and results of the mobilization process; data collection was arranged in order to evaluate the presence or absence of GITMO criteria for pPM.

Statistics

The relevance of the candidate predictive factors was evaluated using univariate logistic regression for the outcome variable of pPM. Subsequently, multiple logistic regression with backward variable selection was performed to identify independent predictive factors. Explored variables are reported in table 1. WBC and absolute neutrophil counts were analyzed on the log-scale because of highly skewed distributions. Continuous parameters were not categorized a priori because this would have negatively affected the power of the analysis. Values of non-dichotomous categorical variables were transformed in dummy variables for the purpose of the analysis.

The outcome variable was the failure of a mobilization attempt defined according to the GITMO criteria for proven poor mobilizer. To this end, in patients treated with Plerixafor on demand, the data collected reflected the situation after the declaration of failure (i.e declining CD34+ cell count with a peak value <20/mcl or at least 3 aphereses with total collection <2 x 10^6 CD34/kg) and before Plerixafor administration. Conversely, patients treated with upfront Plerixafor had their data collected at the end of the mobilization process, as for all other patients.

To estimate the discriminating power of a chosen model, a receiver operating characteristic (ROC) curve was plotted. The areas under the ROC curves (AUCs) were calculated as previously described[16]. AUC comparisons were performed according to the method described by DeLong et al.[17]

Internal validation was performed applying the refined bootstrap described by Efron.[18] Random data splitting in training and validation sample was not performed because this internal validation procedure reduces the power for both model development and validation and is known to be inferior to bootstrap validation. Bootstrap validation used the AUC as performance index.

Two groups were defined by categorizing the score (linear predictor) of the final logistic regression model. For each cut-off, sensitivity, specificity, PPV, and NPV were calculated as simple proportions with 95% confidence intervals (CI). Likelihood ratios and their CI were calculated as ratios between proportions. The McNemar chi-square test was used to compare sensitivity and specificity between assays among failures and non-failures, respectively.[19] Cutpoint selection was based on clinical criteria: the purpose of the clinical tool for PPM prediction was to identify patients at very high risk for mobilization failure in order to support a practice-changing clinical decision. Therefore we aimed to maximize positive likelihood ratio (LR+) over negative likelihood ratio (LR-), by achieving a +LR value >10.

An explorative simplification of the final model was developed using basic mathematical operations. Spearman's rho was calculated to measure the correlation between the original score and the simplified version.[20]

Sample size calculation was based on AUC for the outcome variable of failed mobilization attempt (PPM): assuming a prevalence of PPM equal to 0.2, data from 845 mobilizations (169 failures) had to be collected to obtain an AUC≥ 0.57 (alpha=0.05 and power=0.8); with different PPM prevalences (0.1-0.5), the total number of mobilization attempts to be collected ranged from 530 to 1600.

Statistical analyses were performed using Stata 12 (Statacorp, College Station, Texas) and MedCalc (MedCalc Software, Ostend, Belgium). Significance level was 0.05 for all analyses.

Results

Patient characteristics

We analyzed data from 1318 mobilization attempts. Disease distribution was the following: 600 (46%) patients were affected by MM, 554 (42%) by NHL and 164 (12%) by HL. Median age at diagnosis was 56 years (range 5-76 years); four patients had less than 14 years at diagnosis but underwent mobilization after this age; 56% of patients were male. Sixty percent of patients had been treated with a single chemotherapy course before mobilization, 31% with 2 courses and 8% with 3 or more. Twelve percent of patients had been subjected to treatments potentially harmful to SCM (fludarabine, lenalidomide, radio-immunoconjugates, melphalan, carmustine); extensive radiotherapy on marrow bearing tissue had been used in 23 patients (1.7%). Before the mobilization attempt, 81% of patients were in partial or complete response; BM biopsy (BMB) was negative in 62% and showed extensive infiltration (≥30% of total cellularity) in 3% of patients. Pre-mobilization BMB was omitted in 199 patients, due to different centers' policies (3 centers did not perform it routinely before mobilization).

Priming strategies involved the use of chemotherapy plus G-CSF in 94% of patients; chemotherapy protocols were quite disease-specific: cyclophosphamide was employed mostly in MM patients, while Ara-C containing regimens were preferred in NHL. Upfront plerixafor was added to the mobilization regimen in 44 patients (3%). Ninety-eight patients (7.4%) started SCM with at least one severe cytopenia (≥ grade 3 anemia, thrombocytopenia or neutropenia).

Overall, 180 patients (13.7%) failed the mobilization attempt, according to GITMO criteria for PPM. Failure resulted exclusively from inadequate CD34+ cells mobilization (peak CD34 count <20/mcl) in 36 cases (20%), from insufficient harvest (total CD34 ≤2 x 10^6/kg) in 17 cases (9.4%) and from both criteria in 127 cases (70.6%). Further basal characteristics are reported in table 1.

Validation of the GITMO criteria

To verify the actual consistency of GITMO consensus, we retrospectively applied the criteria to our cohort of 1318 cases. For each case, a score was generated (pPM-GITMO score) by summing 1 point for each minor criteria and 2 points for each major criteria that were present. The only criterion considered in the original publication that could not be ascertained was BMB cellularity before mobilization, given the high rate of missing values.

This score ranged from 0 to 7 and the median value was 1. The AUC relative to the outcome of proven poor mobilizer was 0.673 (95%CI: 0.627-0.719, Fig. 1A). According to the GITMO consensus, the definition of pPM required at least one major criterion or two minor criteria; hence we considered a cut-off equal of greater than 2 for the pPM-GITMO score to be predictive. With this cut-off (Table 2), the sensitivity for the diagnosis of pPM was 53.3% (95%CI: 45.8%-60.8%) and the specificity 73.8% (95%CI: 71.2%-76.3%); LR+ was 2.04 (95%CI: 1.72-2.41). Given the prevalence of proven poor mobilizer observed in our cohort (13.7%), the PPV resulted 24.4% (95%CI: 20.2%-28.9%).

We implemented exploratory analyses to improve the predictive performance of the GITMO-pPM score. Increasing the cut-off to values equal or greater than 3 yielded a significantly lower sensitivity (39.4%) but higher specificity (90.8%); the PPV was 40.3% (95%CI: 33%-48%).

In the GITMO consensus, the splitting into major and minor criteria represented a simplification of the weights derived from AHP; thus we checked whether using the original AHP weights could improve the predictive performance of the GITMO-pPM score. Therefore we generated a score (AHP-pPM score) by summing the relative weight of each criterion as reported in the original publication. This score ranged from 0 to 0.55, had median value of 1 and produced an AUC of 0.679 (95%CI: 0.634-0.725, Fig. 1B). To maximize specificity and LR+, we chose a cut-off equal of greater than 0.21 (Table 2), yielding a sensitivity of 33.3% (95%CI: 26.5%-40.7%) and a specificity of 93.1% (95%CI: 91.4%-94.5%); LR+ was 4.8 (95%CI: 3.57-6.46); PPV was 43.2%% (95%CI: 34.8%-51.8%).

Predictive factors for poor stem cell mobilization

Sex, BMB at diagnosis and previous radiotherapy (local or extensive) did not show predictive relevance for mobilization failure in univariate analyses (Table 3). Among non-dichotomous categorical variables, BMB before mobilization had a significant protective effect if pathologic infiltration was absent, while it favored failure when disease infiltration reached 30% or more. NHL was strongly associated with failure, while HL was the opposite, and MM was non-significant; among priming strategies, use of G-CSF alone had strong impact on failure, while other chemotherapy regimens were not significant. Increasing age, number of full chemotherapy courses, previous use of fludarabine, lenalidomide, melphalan and carmustine, previous mobilization failure, refractory disease, and lower CBC values before mobilization, all had significant negative impact on the main outcome; upfront plerixafor use was instead associated with a reduced probability of failure.

Predicted Poor Mobilizer (pPM) score

Hodgkin's lymphoma, refractory disease, absent pathologic infiltration at pre-mobilization BMB lost predictive relevance when evaluated in multivariate analysis. Continuous variables were categorized to help their potential application in clinical practice. For the same reason, the 4 variables reporting for previous use of fludarabine, lenalidomide, melphan and carmustine were merged in one binary variable encoding for patients undergoing at least one of those treatment at risk. In the final model (Table 4), the following variables were identified as independent predictive factors for mobilization failure: increasing age (from ≤45 years to 46-60 years and to >60 years), diagnosis of NHL, disease infiltration ≥ 30% at the pre-mobilization BMB, previous mobilization failure, increasing number of full chemotherapy courses, previous treatment at risk (fludarabine, lenalidomide, melphan or carmustine), reduced hemoglobin (from >130 g/l to 80-130 g/l to less than 80 g/l), low WBC count (<5 x 10^9/L), low Plt count (<170 x 10^9/L), use of G-CSF alone as a priming strategy and not providing upfront Plerixafor. The predicted poor mobilizer score (pPM score) was calculated as shown in Table 5.

Predicted poor mobilizer score ranged from 2.46 to 12.82, had median value of 5.78 and produced an AUC of 0.801 (Fig. 1C; 95%CI: 0.764-0.838, Fig. 1C). We chose a cut-off >7.862 (Table 2), yielding a specificity of 97.4% (95%CI: 96.3%-98.2%) and a sensitivity of 32.8% (95%CI: 26%-40.2%); LR+ was 12.43 (95%CI: 8.25-18.74), PPV was 66.3% (95%CI: 55.5%-76%).

The probability of mobilization failure according to the pPM score can be calculated as:

$$Probability = \frac{e^{(pPMscore - 8.245)}}{e^{(pPMscore - 8.245)} + 1}$$

The internal validation procedure correcting for overoptimism by bootstrap showed stability of predictive performance measured with AUC values (Table 6).

Simplified predicted Poor Mobilizer score

The classification according to the pPM-score involves some mathematical operations best performed using an electronic calculator. To make the score most practicable, we exploratively simplified it by rounding the weights of each factor to multiple of 0.5 points. This score was calculated as shown in Table 5.

The simplified version of the pPM score was highly correlated with the original one (Spearman's rho =0.983, p <0.0001). Simplified pPM score ranged from 2 to 10, had median value of 4.5 and produced an AUC of 0.795 (Fig. 1D; 95%CI: 0.757-0.833, Fig. 1C). We chose a cut-off ≥ 6.5 (Table 2), yielding a specificity of 98% (95%CI: 97%-98.7%) and a sensitivity of 31.7% (95%CI: 24.9%-39%); LR+ was 15.7 (95%CI: 9.9-24.8); PPV was 71.3% (95%CI: 60%-80.8%).

Score comparison

The AUC of the 4 different scores were compared: GITMO-pPM and AHP-pPM score had both a significantly inferior AUC than pPM score and simplified pPM-score (p<0.0001 for all comparisons). There were no significant differences between GITMO-pPM and AHP-pPM score (p =0.40) and between pPM score and simplified pPM-score (p =0.08). Detailed results are reported in Tab S2.

We next compared the sensitivity and specificity of the different scores according to the chosen cut-offs: GITMO-pPM score with cut-off ≥ 2 had the best sensitivity compared to all alternatives; simplified pPM score with cut-off ≥ 6.5 had the best specificity with respect to all other models and cut-offs. Detailed results are reported in Tab S2.

Discussion

In this retrospective study, we collected a representative sample of mobilization outcomes in the plerixafor era in MM and lymphoma patients. The analysis of this large database aimed: 1-to validate published GITMO criteria for pPM (which were developed by AHP consensus method) on strong clinical data; 2-to improve the predictive ability of these criteria, by adding new variables and refining weights of already present criteria. The ultimate objective was to develop a standardized clinical tool able to identify "a priori" those patients at very high risk of failure, before starting the mobilization procedure, in order to drive a practice-changing clinical decision. Performance measures for prediction of mobilization failure were derived for 4 different models: (1) based on the original GITMO criteria (GITMO-pPM score); (2) using original AHP weights of GITMO criteria (AHP-pPM score); (3) a new model derived through multivariate regression analysis (pPM score); (4) a simplified version of this new model (simplified pPM score). The original GITMO criteria had modest performance measured by AUC (0.67); when applied with the proposed cut-off for pPM, it had limited sensitivity (53%) and modest specificity (74%) and use of original AHP weights did not improve their predictive performance. The new model (pPM-score) had far better AUC (0.80); its simplified version (ranging from 2 to 10) was categorized using a cut-off to maximize specificity: indeed in our sample, a high proportion of patients with simplified pPM-score >6.5 failed the mobilization (PPV =71%). Simplified pPM-score, combining unmodifiable patient-related factors with clinical choice-dependent variables, can be easily simulated before starting SCM, therefore supporting patient-tailored mobilization strategies..

Today, the first key decision in scheduling a first-line SCM regimen is the choice between a chemo-mobilization or a cytokine-only strategy. The second crucial stage is the dynamic identification of those patients, during SCM, in whom the addition of just-in-time plerixafor could be useful and cost-effective. To this end, different algorithms have been proposed:[20,21,22] all of

them include PB CD34+ cell count, the most reliable parameter to trigger plerixafor administration.[23,24] Other parameters proposed include WBC and platelet counts as surrogates of hematopoietic recovery, collection target dose and first day of apheresis yield.[9] Nevertheless, such algorithms present several limits. First of all, circulating CD34+ threshold values used to trigger plerixafor administration present a significant variability between different studies, ranging from 7^{27} to $10^{9,15}$ or 20/mcl.[25] Secondly, most of those algorithms were not validated outside the institution they were developed, making problematic their application to other centers, as significant differences exist in facilities, staff, skills and procedures. In addition, many algorithms leave unresolved a "gray zone" with intermediate values of PB CD34+ (i.e. 10-20/mcl), where no recommendations are drawn and a "case by case" approach is suggested. The EBMT recommendations[26] recognized this window of uncertainty and proposed to fill the gap with a clinical decision taking into account risk factors for poor mobilization. Although these recommendations acknowledged first the role of patient and disease-related risk factors in the decision-making of mobilization, the choice was left to individual discretion. Recently published US recommendations[27] suggest as well to tailor the mobilization plan according to patient and disease characteristics; in case of MM, the authors suggest chemo-mobilization (instead of steady-state strategy) for patients previously treated with lenalidomide or melphalan, or having received more than 1 previous line of therapy. Similarly, for patient with lymphoma, the authors recommend to limit steady-state mobilization to patients "at low risk for mobilization failure"; once again, an explicit and reproducible clarification of the "risk of failure" is missing.

In 2012 a GITMO panel of experts proposed definitions for PM, recognizing two clearly different categories: the one of proven PM, which referred to a completed process merely requiring uniform and detailed characterization; the other one of predicted PM, i.e. a new classification of

patients expected to be at higher risk of failure for future mobilizations. Although application of AHP methodology to complex issues demonstrated excellent results even when consistent data were unavailable[28], the adoption of a predictive model in a clinical setting requires nonetheless validation on real life patients. To verify the actual consistency of GITMO consensus, we retrospectively applied the criteria to our cohort of 1318 cases and measured their predictive performance with the AUC. Although no single measure of diagnostic accuracy fully captures the clinical value of a test, AUC is considered a valuable estimate of the global discriminative power, being independent from the chosen cut-off and from the disease prevalence.[29] Although the threshold to reach our predefined endpoint was set low (0.57), the obtained value (0.67) is considered indicative of sufficient diagnostic accuracy.

To gain a significant improvement over the GITMO criteria, we elaborated a new score based on data collected in our large database, evaluating all variables originally considered by the GITMO consensus and new ones: several risk factors previously identified by the Consensus were confirmed as relevant, such as stem cell poisons (e.g. lenalidomide and fludarabine). Instead, the role of extensive radiotherapy (previously considered major criterion) did not emerge as statistically significant, probably due to the very low number of patients who actually received it. Finally, a relevant statistical weight emerged for blood counts of all lineages, which adds to confirmation of other factors already identified by the Consensus (neoplastic BM involvement and previous chemotherapy burden), to allow detailed characterization of the BM functional reserve with simple parameters.

The pPM score undoubtedly improved the diagnostic accuracy with respect to GITMO criteria. However, improvements in test accuracy will not benefit patients unless they lead to changes in patient management.[30] To reach this goal, a clinical test should be easily applied and interpreted.

The pPM score contains predictors that are known before starting SCM: some pertain to patients' history, others to procedures routinely performed in the clinical practice (CBC, BMB), others to the mobilization planning. Algorithms based on PB CD34+ cell count only allow a late-stage clinical decision, when the mobilization process is already close to the end and only limited action (addition of just-in-time plerixafor) is possible. A finer and earlier planning would add a significantly wider range of possibilities to improve mobilization outcomes (table 7). To enhance feasibility of pPM-score, we created a simplified version, easily computable without an electronic calculator. Furthermore, we decided to dichotomize simplified pPM-score, aiming to make it a "ruling-in" diagnostic test. To this end we chose to maximize the LR+, albeit preserving a sensitivity ≥30%. Many useful properties make LR+ suitable to this scope[31]: LR+ is independent of the disease prevalence in the examined group, making it immediately applicable to other clinical settings; LR+ is considered the best indicator for ruling-in diagnosis: the higher the LR+, the greater is the shift of the probability of disease. Good diagnostic tests have LR+ >10 and their positive result has a significant contribution to the diagnosis. The simplified pPM-score with a cut-off ≥6.5 had a LR+ of 15.7: in our sample, this means that positive patients have their probability of failure increased from 13.7% to 71.3%, therefore requiring alternative strategies to avoid highly likely failure. One option is tailoring the priming strategy: the eternal dispute between steady-state and chemo-mobilization would be resolved if we could appropriately personalize mobilization strategies. Our results clearly support the use of chemo-mobilization in pPM, confirming a growing body of evidence.[32] Second, we provide a strong suggestion for upfront use of Plerixafor in pPM. Interestingly, very recently Yuan and colleagues[33] reported the results of the mobilization policy implemented in their center in Duarte, California. They propose the administration of upfront plerixafor in patients defined as "predicted poor mobilizers" according to criteria similar to ours,

and in MM patients candidates for tandem auto-SCT. Alternatively, the pPM-score can be integrated in algorithms based on PB CD34+ cell count, to resolve their "gray zones", thus justifying earlier addition of Plerixafor. Finally, recognizing in advance pPMs will enable a special surveillance on them during the mobilization process, allowing for several technical optimizations such as use of large volume apheresis or starting apheresis with lower CD34+ thresholds (table 7).

Maximization of the LR+ implied an obvious reduction of the sensitivity (32% using cut-off ≥6.5): this important limitation should be taken into account if simplified pPM-score is used for clinical decisions. Negative patients should still be considered at risk for mobilization failure. In this group, we suggest careful monitoring of mobilization kinetics and application of algorithms for "just-in-time" Plerixafor to rescue additional patients from mobilization failure.

Another important issue to be considered when applying the pPM-score is that minimum dose of CD34+ is not equivalent to target dose: in our analysis, failure was defined as collection of less than 2 $x10^6$ CD34+/kg. However, in the clinical practice, definition of failure should be related to patient's goals: as an example, collection of 3 x 10^6 CD34+/kg is clearly unsatisfactory if a double auto-SCT is planned.

In conclusion, we have built on real large representative data a score to "rule in" patients at very high risk for PM before starting mobilization, allowing changes in clinical management, to avoid highly likely mobilization failure. To achieve the highest possible power from the available data, we performed internal validation by bootstrap,[34] thus confirming a high stability of the developed predictive model. Nevertheless, an external validation on an independent data set is still required to allow a broad application of this clinical tool. Finally, given the retrospective and observational nature of this study, it should be reminded that changes of SCM strategies which may be

suggested by pPM-score application (table 7), although reasonable, warrant to be tested in a prospective interventional trial to demonstrate clinical effectiveness.

Conflicts of interest:

The authors do not have any conflicts of interest.

Acknowledgments:

The authors are very grateful to Giuseppe Ausoni, Paola Brambilla, Saveria Capria, Gloria Margiotta Casaluci, Michele Cimminiello, Annarita Conconi, Carmela Cuomo, Katia Codeluppi, Mario Delia, Roberta Distefano, Annalisa Di Marco, Luca Facchini, Salvatore Gattillo, Maria Gozzer, Svitlana Gumenyuk, Francesco Marchesi, Giovanna Meloni, Angela Melpignano, Luca Nassi, Domenico Pastore, Giuseppe Pietrantuono, Michele Pizzuti, Giovanni Quarta, Azzurra Anna Romeo, Federica Sorà, Andrea Spadaro and Stefania Trinca for their contribution to this study.

Author contributions

JO performed the statistical analysis and wrote the manuscript; AO contributed to study design, interpreted the results, contributed to manuscript writing, reviewed and approved the manuscript; EDN contributed to study design and to the statistical analysis; FS contributed to data collection and interpretation of the results, approved and edited the manuscript; IA, MC, MP, PP, PEP contributed to data collection and interpretation of the results; PC, LF, GG, LN, SS, NP, MM, TM, MP, FZ, FC, SM, AM, PM, SC, FM, KC, GM, FL, GS, DP, GM contributed to contributed to patient care and data collection.

Figure Legends

Figure 1: Area Under the Receiving Operating Characteristic (ROC) Curve (AUC) for the outcome of proven poor mobilizer and the 4 scores generated (A: GITMO-PPM score; B: AHP-PPM score; C: PPM score; D: simplified PPM score).

REFERENCES

[1]Passweg JR, Baldomero H, Bader P, Bonini C, Cesaro S, Dreger P, et al. Hematopoietic stem cell transplantation in Europe 2014: more than 40 000 transplants annually. *Bone Marrow Transplant* 2016; 51: 786-92.

[2]Weaver CH, Hazelton B, Birch R, Palmer P, Allen C, Schwartzberg L et al. An analysis of engraftment kinetics as a function of the CD34 content of peripheral blood progenitor cell collections in 692 patients after the administration of myeloablative chemotherapy. *Blood* 1995; 86: 3961–3969.

[3]Hubel K, Fresen MM, Apperley JF, Basak GW, Douglas KW, Gabriel IH et al. European data on stem cell mobilization with plerixafor in non-Hodgkin's lymphoma, Hodgkin's lymphoma and multiple myeloma patients. A subgroup analysis of the European Consortium of stem cell mobilization. *Bone Marrow Transplant* 2012; 47: 1046-50.

[4]Perseghin P, Terruzzi E, Dassi M, Baldini V, Parma M, Coluccia P et al. Management of poor peripheral blood stem cell mobilization: incidence, predictive factors, alternative strategies and outcome. A retrospective analysis on 2177 patients from three major Italian institutions. *Transfus Apher Sci* 2009; 41: 33–37.

[5]Pusic I, Jiang SY, Landua S, Uy GL, Rettig MP, Cashen AF, et al. Impact of mobilization and remobilization strategies on achieving sufficient stem cell yields for autologous transplantation. *Biol Blood Marrow Transplant* 2008; 14: 1045–1056.

[6]Milone G, Martino M, Spadaro A, Leotta S, Di Marco A, Scalzulli P et al. Plerixafor on demand combined with chemotherapy and granulocyte colony-stimulating factor: significant improvement in

peripheral blood stem cells mobilization and harvest without increase in costs. *Br J Haematol* 2013; 164: 113-23

[7]Farina L, Guidetti A, Spina F, Roncari L, Longoni P, Ravagnani F et al. Plerixafor 'on demand': results of a strategy based on peripheral blood CD341 cells in lymphoma patients at first or subsequent mobilization with chemotherapy + G-CSF. *Bone Marrow Transplant* 2014; 9: 453-5.

[8]Sancho JM, Morgades M, Grifols JR, Juncà J, Guardia R, Vives S et al. Predictive factors for poor peripheral blood stem cell mobilization and peak CD34(+) cell count to guide pre-emptive or immediate rescue mobilization. *Cytotherapy* 2012; 4: 823–829.

[9]Lanza F, Lemoli RM, Olivieri A, Laszlo D, Martino M, Specchia G et al. Factors affecting successful mobilization with plerixafor: an Italian prospective survey in 15 patients with multiple myeloma and lymphoma. *Transfusion* 2013; 54: 31–339.

[10]Sinha S, Gertz MA, Lacy MQ, Dispenzieri A, Hayman SR, Buadi FK et al. Majority of patients receiving initial therapy with lenalidomide-based regimens can be successfully mobilized with appropriate mobilization strategies. *Leukemia* 2012;

26: 1119–1122.

[11]Giralt S, Stadtmauer EA, Harousseau JL, Palumbo A, Bensinger W, Comenzo RL, et al. International Myeloma Working Group (IMWG) consensus statement and guidelines regarding the current status of stem cell collection and high-dose therapy for multiple myeloma and the role of plerixafor (AMD 3100). *Leukemia* 2009; 23: 1904–12.

[12]Attolico I, Pavone V, Ostuni A, Rossini B, Musso M, Crescimanno A et al. Plerixafor added to chemotherapy plus G-CSF is safe and allows adequate PBSC collection in predicted poor mobilizer patients with multiple myeloma or lymphoma. *Biol Blood Marrow Transplant* 2012; 18: 241– 249.

[13]Olivieri A, Marchetti M, Lemoli R, Tarella C, Iacone A, Lanza F et al. Proposed definition of 'poor mobilizer' in lymphoma and multiple myeloma: an analytic hierarchy process by ad hoc working group Gruppo Italiano Trapianto di Midollo Osseo. *Bone Marrow Transplant* 2012; 47: 342–351.

[14]Abhyankar S, DeJarnette S, Aljitawi O, Ganguly S, Merkel D, McGuirk J. Risk-based approach to optimize autologous hematopoietic stem cell (HSC) collection with the use of plerixafor. *Bone Marrow Transplant* 2012; 47:83–487.

[15]Costa LJ, Alexander ET, Hogan KR, Schaub C, Fouts TV, Stuart RK. Development and validation of a decision-making algorithm to guide the use of plerixafor for autologous hematopoietic stem cell mobilization. *Bone Marrow Transplant* 2011; 6: 64–69.

[16] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29-36.

[17] DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 44: 837–845.

[18]Efron B, Tibshirani RJ. An Introduction to the Bootstrap. Vol. 57: Chapman & Hall/CRC; 1994.

[19]Trajman A, Luiz RR. McNemar chi2 test revisited: comparing sensitivity and specificity of diagnostic examinations. *Scand J Clin Lab Invest*. 2008;68(1):77-80.

[20]Hoster E, Dreyling M, Klapper W, Gisselbrecht C, van Hoof A, Kluin-Nelemans HC, et al. A new prognostic index (MIPI) for patients with advanced-stage mantle cell lymphoma. *Blood* 2008; 111: 558-65.

[21]Micallef IN, Sinha S, Gastineau DA, Wolf R, Inwards DJ, Gertz MA et al. Cost-effectiveness analysis of a risk-adapted algorithm for plerixafor use in autologous peripheral blood stem cell mobilization. *Biol Blood Marrow Transplant*. 2013;19: 87-93.

[22]Horwitz ME, Chute JP, Gasparetto C, Long GD, McDonald C, Morris A et al. Preemptive dosing of plerixafor given to poor stem cell mobilizers on day 5 of G-CSF administration. *Bone Marrow Transplant* 2012; 47: 1051-5.

[23]Gambell P, Herbert K, Dickinson M, Stokes K, Bressel M, Wall D et al. Peripheral blood CD34+cell enumeration as a predictor of apheresis yield: an analysis of over 1000 collections. *Biol Blood Marrow Transplant*. 2012;18:763-772.

[24]Pierelli L, Perseghin P, Marchetti M, Accorsi P, Fanin R, Messina C, et al. Best practice for peripheral blood progenitor cell mobilization and collection in adults and children: results of a Società Italiana Di Emaferesi e Manipolazione Cellulare (SIDEM) and Gruppo Italiano Trapianto Midollo Osseo (GITMO) consensus process. *Transfusion* 2012; 52: 893-905.

[25] Chow E, Rao KV, Wood WA, Covington D, Armistead PM, Coghill J, et al. Effectiveness of an algorithm-based approach to the utilization of plerixafor in patients undergoing chemotherapy-based stem cell mobilization. *Biol Blood Marrow Transplant*. 2014; 20: 1064-8.

[26]Mohty M, Hübel K, Kröger N, Aljurf M, Apperley J, Basak GW, et al. Autologous haematopoietic stem cell mobilisation in multiple myeloma and lymphoma patients: a position statement from the European Group for Blood and Marrow Transplantation. *Bone Marrow Transplant* 2014; 49: 865-72.

[27]Giralt S, Costa L, Schriber J, Dipersio J, Maziarz R, McCarty J, et al. Optimizing autologous stem cell mobilization strategies to improve patient outcomes: consensus guidelines and recommendations. *Biol Blood Marrow Transplant* 2014; 20: 295-308.

[28]Whitaker R. Validation examples of the analytic hierarchy process and analytic network process. *Mathematical and Computer Modelling* 2007; 46: 840-859.

[29]Šimundić AM. Measures of Diagnostic Accuracy: Basic Definitions. EJIFCC. 2009;19(4):203-11.

[30] Ferrante di Ruffano L, Hyde C, McCaffery KJ, Bossuyt PM, Deeks JJ. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. *BMJ* 2012; 344: e686

[31] Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. *Lancet* 2005; 365: 1500-5.

[32] Olivieri A, Saraceni F. Mobilization policy in multiple myeloma: minimum target or law of redundancy? Two different approaches by the two sides of the Atlantic Ocean. *Bone Marrow Transplant* 2016; 51: 348-50.

[33] Yuan S, Wang S. How do we mobilize and collect autologous peripheral blood stem cells? *Transfusion* 2017; 57: 13-23.

[34]Steyerberg EW, Harrell FE, Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001; 54: 774-781

Figure 1: Area Under the Receiving Operating Characteristic (ROC) Curve (AUC) for the outcome of proven poor mobilizer and the 4 scores generated (A: GITMO-PPM score; B: AHP-PPM score; C: PPM score; D: simplified PPM score).

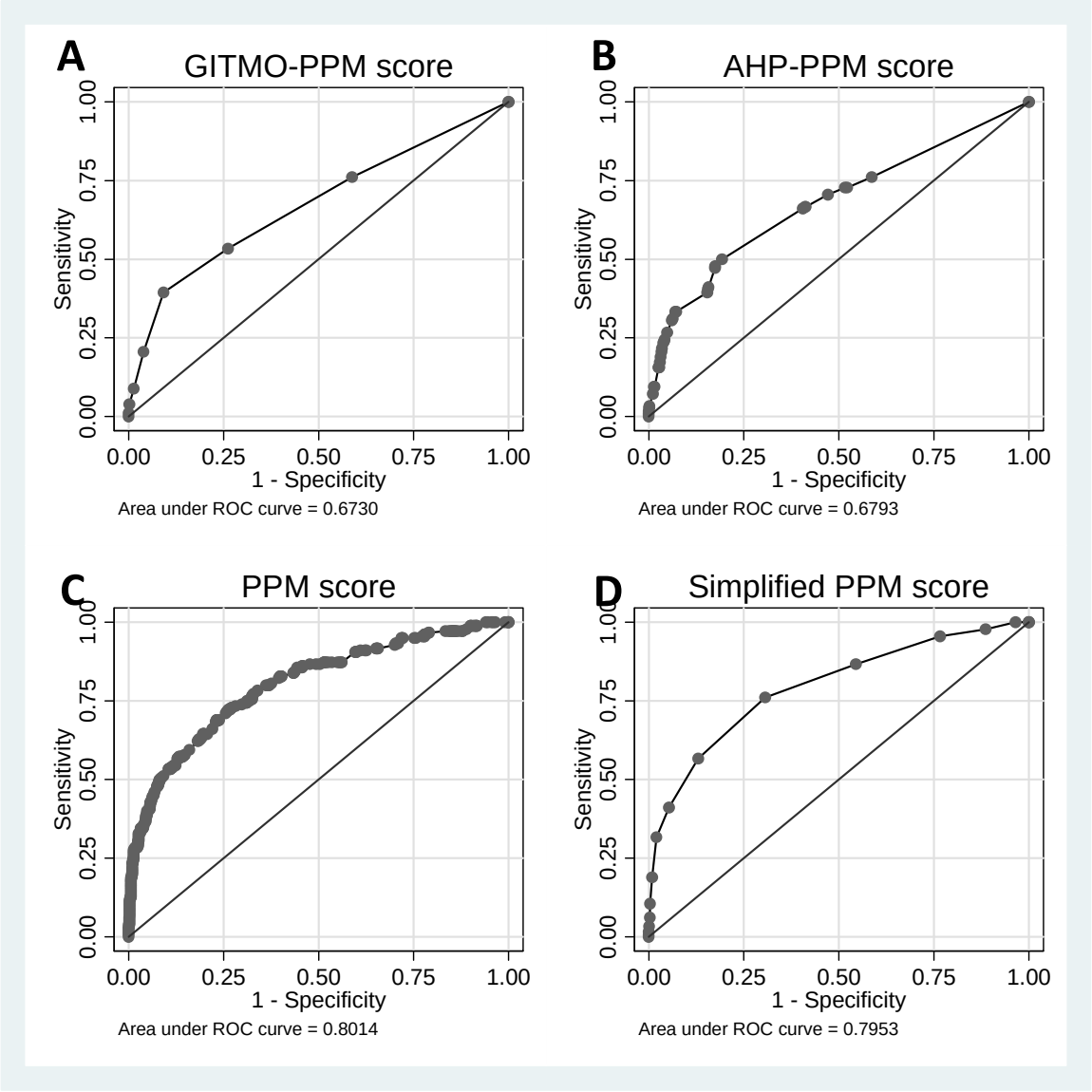

Table I: Basal characteristics

BASAL CHARACTERISTICS	ALL PATIENTS	MM	NHL	HL
Age at diagnosis, median (range)	55.6 (4.6 - 76.5)	59 (18 - 74)	54 (5 - 76)	37 (5 - 76)
≤45	337 (26%)	57 (10%)	166 (30%)	114 (70%)
45-60	571 (43%)	293 (49%)	246 (44%)	32 (20%)
> 60	410 (31%)	250 (42%)	142 (26%)	18 (11%)
Sex				
Male	753 (57%)	321 (54%)	346 (62%)	86 (52%)
Female	565 (43%)	279 (47%)	208 (38%)	78 (48%)
Disease				
Multiple Myeloma	600 (46%)			
Non-Hodgkin's Lymphoma	554 (42%)			
Hodgkin's Lymphoma	164 (12%)			
Bone marrow infiltration at diagnosis				
Absent	499 (38%)	18 (3%)	332 (60%)	149 (91%)
Present: < 30%	306 (23%)	174 (29%)	121 (22%)	11 (7%)
Present: ≥ 30%	511 (39%)	407 (68%)	100 (18%)	4 (2%)
Unknown	2 (0.2%)	1 (0.2%)	1 (0.2%)	0 (0%)
PREVIOUS TREATMENTS				
Number of chemotherapy courses - median (range)	1 (1 - 6)	1 (1 - 6)	2 (1 - 6)	2 (1 - 5)
1	790 (60%)	489 (82%)	259 (47%)	42 (26%)
2	413 (31%)	79 (13%)	230 (42%)	104 (63%)
3	93 (7%)	21 (4%)	57 (10%)	15 (9%)
≥4	22 (2%)	11 (2%)	8 (1%)	3 (2%)
Use of myelotoxic agents (at least one)	153 (12%)	133 (22%)	15 (3%)	5 (3%)
Fludarabine	12 (1%)	1 (0.2%)	8 (1%)	3 (2%)
Lenalidomide (≤ 4cycles / > 4 cycles)	114 (9%) / 7 (1%)	113 (19%) / 7 (1%)	1 (0.2%) / 0 (0%)	0 (0%) / 0 (0%)
Radioimmunoconjugates	1 (0.1%)	1 (0.2%)	0 (0%)	0 (0%)
Melphalan	27 (2%)	20 (3%)	5 (1%)	2 (1%)
BCNU	9 (1%)	2 (0.3%)	5 (1%)	2 (1%)
Radiotherapy (limited / extensive)	122 (9%) / 23 (2%)	42 (7%) / 21 (4%)	37 (7%) / 6 (1%)	43 (26%) / 5 (3%)
Bone marrow infiltration before mobilization				
Absent	821 (62%)	201 (34%)	471 (85%)	149 (91%)

Present: < 30%	263 (20%)	231 (39%)	26 (5%)	6 (4%)
Present: ≥ 30%	35 (3%)	30 (5%)	5 (1%)	0 (0%)
Unknown	199 (15%)	138 (23%)	52 (9%)	9 (5%)
Disease status at mobilization				
Remission (complete or partial)	1066 (81%)	535 (89%)	426 (77%)	105 (64%)
Refractory	242 (18%)	63 (11%)	121 (22%)	58 (35%)
Unknown	10 (1%)	2 (0.3%)	7 (1%)	1 (1%)
Failed previous mobilization attempt	94 (7%)	37 (6%)	51 (9%)	6 (4%)
MOBILIZATION				
Blood count values before starting mobilization				
Hemoglobin (g/dl) - median (range)	11.8 (7.2 - 19.8)	12.2 (7.2 - 18.8)	11.3 (7.2 - 17.3)	11.6 (7.9 - 16.2)
Leukocytes (x 10^9/L) - median (range)	5.2 (0 - 426)	5.3 (0.3 - 42.6)	4.9 (0 - 58.27)	5.8 (0.9 - 22.8)
Neutrophils (x 10^9/L) - median (range)	3.2 (0 - 282)	3.1 (0.1 - 28.2)	3.1 (0 - 45.2)	3.9 (0.1 - 18.7)
Platelets (x 10^9/L) - median (range)	223 (6 - 1167)	230 (6 - 665)	202 (6 - 1167)	239 (7 - 601)
Mobilization regimen				
High dose CTX (2-7 g/mq) + G-CSF	650 (49%)	499 (83%)	131 (24%)	20 (12%)
DHAP + G-CSF	126 (10%)	6 (1%)	101 (18%)	19 (12%)
IEV + G-CSF	70 (5%)	0 (0%)	21 (4%)	49 (30%)
High dose Ara-C + G-CSF	107 (8%)	4 (1%)	100 (18%)	3 (2%)
Other chemotherapy regimen + G-CSF	292 (22%)	43 (7%)	177 (32%)	72 (44%)
G-CSF alone	73 (6%)	48 (8%)	24 (4%)	1 (1%)
Dose of G-CSF (μg/kg)				
5	904 (69%)	332 (55%)	441 (80%)	131 (80%)
10	413 (31%)	267 (45%)	113 (20%)	33 (20%)
15	1 (0.1%)	1 (0.2%)	0 (0%)	0 (0%)
Plerixafor administered (upfront)	44 (3%)	18 (3%)	24 (4%)	2 (1%)
Peak CD34+ value (cells/mcl) - median (range)	85 (0 - 1942)	89 (0 - 971)	76 (0 - 1942)	107 (0 - 1231)
<5	61 (5%)	20 (3%)	36 (6%)	5 (3%)
<20	163 (12%)	66 (11%)	89 (16%)	8 (5%)
Number of aphereses - median (range)	1 (0 - 6)	2 (0 - 6)	1 (0 - 5)	1 (0 - 5)
>3	47 (4%)	32 (5%)	10 (2%)	5 (3%)
Total harvest (x 10^6 CD34+/kg) - median (range)	8.9 (0 - 63.5)	10.1 (0 - 44)	7.6 (0 - 63.5)	9.1 (0 - 47.9)
<1	118 (9%)	47 (8%)	64 (12%)	7 (4%)

<2	144 (11%)	59 (10%)	76 (14%)	9 (5%)
2 – 5	204 (15%)	87 (15%)	93 (17%)	24 (15%)
>5	970 (74%)	454 (76%)	385 (69%)	131 (80%)
Failed mobilization	180 (14%)	75 (13%)	95 (17%)	10 (6%)
Due to low CD34+ peak count	163 (12%)	66 (11%)	89 (16%)	8 (5%)
Due to insufficient harvest	144 (11%)	56 (9%)	76 (14%)	9 (5%)
Due to both above criteria	127 (10%)	50 (8%)	70 (13%)	7 (4%)

Table II: Measures of sensitivity, specificity, positive (LR+) and negative (LR-) likelihood ratio, diagnostic odds ratio, positive (PPV) and negative (NPV) predictive values for selected cut-offs of the 4 scores generated to predict mobilization failure.

	GITMO-PPM score		AHP-PPM score	PPM score		Simplified PPM score	
Cut-off	≥ 2	≥ 3	≥ 0.21	> 7.48	> 7.862	≥ 6	≥ 6.5
Sensitivity (SE)	0.533 (0.458 - 0.608)	0.394 (0.323 - 0.47)	0.333 (0.265 - 0.407)	0.4 (0.328 - 0.476)	0.328 (0.26 - 0.402)	0.411 (0.338 - 0.487)	0.317 (0.249 - 0.39)
Specificity (SP)	0.738 (0.712 - 0.763)	0.908 (0.889 - 0.924)	0.931 (0.914 - 0.945)	0.952 (0.938 - 0.963)	0.974 (0.963 - 0.982)	0.947 (0.933 - 0.96)	0.98 (0.97 - 0.987)
Positive Likelihood Ratio (LR+)	2.04 (1.72 - 2.41)	4.28 (3.31 - 5.53)	4.8 (3.57 - 6.46)	8.28 (6.05 - 11.33)	12.43 (8.25 - 18.74)	7.8 (5.76 - 10.55)	15.67 (9.91 - 24.77)
Negative Likelihood Ratio (LR-)	0.63 (0.54 - 0.74)	0.67 (0.59 - 0.75)	0.72 (0.65 - 0.8)	0.63 (0.56 - 0.71)	0.69 (0.62 - 0.76)	0.62 (0.55 - 0.7)	0.7 (0.63 - 0.77)
Diagnostic Odds Ratio	3.22 (2.34 - 4.44)	6.41 (4.47 - 9.18)	6.7 (4.57 - 9.84)	13.13 (8.79 - 19.62)	18.01 (11.2 - 28.96)	12.54 (8.46 - 18.59)	22.47 (13.42 - 37.58)
Positive predictive value (PPV)	0.244 (0.202 - 0.289)	0.403 (0.33 - 0.48)	0.432 (0.348 - 0.518)	0.567 (0.476 - 0.655)	0.663 (0.555 - 0.76)	0.552 (0.464 - 0.638)	0.713 (0.6 - 0.808)
Negative predictive value (NPV)	0.909 (0.889 - 0.927)	0.905 (0.886 - 0.921)	0.898 (0.88 - 0.915)	0.909 (0.892 - 0.925)	0.902 (0.884 - 0.918)	0.91 (0.893 - 0.926)	0.901 (0.883 - 0.917)

Table III: Association to mobilization failure according to univariate logistic regression

Candidate predictive factor	Odds ratio (95% CI)	Probability (Wald test)
Age (years)	1.01 (1 - 1.03)	0.033
Sex (female)	1.23 (0.89 - 1.68)	0.204
Disease		
MM	0.83 (0.61 - 1.15)	0.264
NHL	1.65 (1.21 - 2.27)	0.002
HL	0.38 (0.19 - 0.73)	0.004
BMB at diagnosis		
Absent	0.99 (0.72 - 1.37)	0.967
<30%	0.71 (0.48 - 1.06)	0.094
≥ 30%	1.27 (0.93 - 1.75)	0.135
Number of chemotherapy courses	2.03 (1.69 - 2.45)	<0.001
Previous use of BCNU	5.15 (1.37 - 19.35)	0.015
Previous use of Fludarabine	4.61 (1.45 - 14.69)	0.010
Previous use of Melphalan	7.29 (3.37 - 15.79)	<0.001
Previous use of Lenalidomide		
Absent	0.43 (0.28 - 0.68)	<0.001
≤ 4 cycles	2.25 (1.42 - 3.57)	0.001
>4 cycles	2.55 (0.49 - 13.22)	0.266
At least one treatment at risk	2.93 (1.98 - 4.35)	<0.001
Previous radiotherapy		
Absent	0.89 (0.55 - 1.43)	0.623
Limited	1.1 (0.65 - 1.87)	0.711
Extensive (on marrow bearing tissue)	1.18 (0.45 - 3.09)	0.743
Previous mobilization failure	6.36 (4.08 - 9.9)	<0.001
Disease remission (CR or PR) before mobilization	1.96 (1.37 - 2.81)	<0.001
Pre-mobilization BMB		
Absent	0.74 (0.54 - 1.02)	0.066
<30%	1.04 (0.71 - 1.54)	0.828
≥ 30%	3.02 (1.45 - 6.28)	0.003
Notdone	1.2 (0.79 - 1.83)	0.392
CBC before mobilization		
Hemoglobin (10 g/L)	0.81 (0.74 - 0.9)	<0.001
Leukocytes (10-fold)	0.31 (0.18 - 0.55)	<0.001
Neutrophils (10-fold)	0.39 (0.24 - 0.64)	<0.001
Platelets (1 x 10^9/L)	0.996 (0.994 – 0.997)	<0.001
Priming strategy		
CTX 3-7 g/m^2 + G-CSF	0.78 (0.57 - 1.07)	0.118
DHAP or DHAOx + G-CSF	0.58 (0.31 - 1.1)	0.094
IEV + G-CSF	0.58 (0.25 - 1.36)	0.208
High-dose Ara-C + G-CSF	0.95 (0.53 - 1.7)	0.857
Other chemotherapy + G-CSF	0.97 (0.66 - 1.42)	0.865
G-CSF alone	5.43 (3.31 - 8.9)	<0.001

Type of G-CSF		
Lenograstim	0.9 (0.65 - 1.25)	0.536
Filgrastim	1.23 (0.9 - 1.69)	0.197
Pegfilgrastim	1.47 (0.41 - 5.2)	0.553
Biosimilar	0.79 (0.18 - 3.46)	0.752
Missing data	0.63 (0.32 - 1.23)	0.179
Double G-CSF dose (vs standard)	1.02 (0.73 - 1.43)	0.918
Upfront plerixafor	2.47 (1.25 - 4.89)	0.010

Table IV: Independent predictive factors for mobilization failure identified by backward variable selection with multiple logistic regression on significance level 0.1 for the Wald statistic

Predictive factor	β	Odds ratio (95% CI)	Probability (Wald test)
Age class (46-60 years = 1; > 60 years = 2)	0.3796	1.46 (1.14 - 1.88)	0.003
Diagnosis = NHL	0.5535	1.74 (1.16 - 2.6)	0.007
Disease infiltration ≥ 30% at the pre-mobilization BMB	1.269	3.56 (1.51 - 8.35)	0.004
Number of full chemotherapy courses	0.5888	1.8 (1.43 - 2.27)	<0.001
At least one previous treatment at risk	0.7739	2.17 (1.28 - 3.67)	0.004
Pre-mobilization Hb value class (<80 g/l = 1; 80 – 130 g/l = 2)	1.1165	3.05 (1.72 - 5.42)	<0.001
Pre-mobilization WBC < 5 x 10^9/L	0.7185	2.05 (1.41 - 2.99)	<0.001
Pre mobilization Plt< 170 x 10^9/L	0.5869	1.8 (1.23 - 2.62)	0.002
Priming with G-CSF alone	2.2513	9.5 (4.75 - 19)	<0.001
Upfront Plerixafor not planned	2.7292	15.32 (5.09 - 46.16)	<0.001
Previous mobilization failure	1.9059	6.73 (3.67 - 12.34)	<0.001

Table V. Calculation of *predicted Poor Mobilizer (pPM)* and *simplified predicted Poor Mobilizer (s-PPM)score*

pPM score	Simplified pPM score
0.3796 (if age 46-60 years) + 0.7592 (if age > 60 years) + 0.5535 (if diagnosis NHL) + 1.269 (if disease infiltration ≥ 30% at the pre-mobilization BMB) + 0.5888 x [number of full chemotherapy courses] + 0.77388929 (if one previous treatment at risk) + 1.1165 (if Hb 80 – 130 g/l) + 2.233 (if Hb < 80 g/l) + 0.7185 (if WBC count < 5 x 109/L) + 0.5869 (if Plt < 170 x 109/L) + 2.251 (if priming with G-CSF alone) + 2.7292 (if upfront Plerixafor not planned) + 1.906 (if previous mobilization failure)	0.5 (if age > 60 years) + 0.5 (if diagnosis NHL) + 1 (if disease infiltration ≥ 30% at the pre-mobilization BMB) + 0.5 x [number of full chemotherapy courses] + 0.5 (if one previous treatment at risk) + 1 (if Hb 80 – 130 g/l) + 2 (if Hb < 80 g/l) + 0.5 (if WBC count < 5 x 109/L) + 0.5 (if Plt < 170 x 109/L) + 2 (if priming with G-CSF alone) + 2 (if upfront Plerixafor not planned) + 1.5 (if previous mobilization failure)

Table VI: Bootstrap validation according to Efron et al.

Apparent area under the ROC curve	0.8014
Mean AUC of 10.000 bootstrap samples	0.8066
Mean AUC of 10.000 tests in original database	0.7959
Optimism in apparent performance	0.0107
Optimism-corrected AUC	0.7907

Table 7. The pPM score can be used to tailor SCM strategy from baseline, unmodifiable risk factors: here we describe examples of calculation of pPM score in 10 different clinical scenarios; we also report suggested SCM strategies based on changes of pPM score due to different clinical choices.

Category of patients	Baseline pPM score	Suggested SCM strategy	pPM score	Predicted probability of failure	Other suggestions
Low risk					
MM with ≥ PR after 1st line (without lenalidomide), no cytopenias, even beyond 60y	1	Cytokine-only (+2) No upfront PLX (+2)	5	<15%	Tailor collection according to target dose
HL without significant BM involvement after 2nd line, <60y	1		5	<15%	
NHL without significant BM involvement after 1st line, <60y	1		5	<15%	
Intermediate risk					
NHL without significant BM involvement after 1st salvage treatment, mild cytopenias*, > 60y	3	CHT-based SCM (0) No upfront PLX (+2)	5	12-16%	Careful SCM monitoring with prompt intervention (Plx on demand)
MM with marrow plasmacytosis <30% after 2nd line with lenalidomide, mild cytopenias*, > 60y	3		5	15-18%	
MM with marrow plasmacytosis <30% after 2nd line with lenalidomide, no cytopenias, after failed SCM attempt, > 60y	3.5	Cytokine-only (+2) Upfront PLX (0)	5.5	<30%	
High risk					
NHL (no significant BM involvement) after 1st salvage treatment, after failed SCM attempt, mild cytopenias*, > 60y	4.5	CHT-based SCM (0) No upfront PLX (+2)	6.5	50%	On demand Plx (if not planned upfront) Use of large volume apheresis Start apheresis with lower CD34+ threshold
		CHT-based SCM (0) Upfront PLX (0)	4.5	<10%**	
MM with marrow plasmacytosis >30% after 2nd line with lenalidomide, trilinear cytopenias, > 60y	5	CHT-based SCM (0) No upfront PLX (+2)	7	71%	
		CHT-based SCM (0) Upfront PLX (0)	5	<15%**	
NHL (no significant BM involvement) after 2nd salvage treatment (of which one at risk), trilinear cytopenias, > 60y	5	CHT-based SCM (0) No upfront PLX (+2)	7	68%	
		CHT-based SCM (0) Upfront PLX (0)	5	<15%**	

MM: Multiple Myeloma; NHL: Non-Hodgkin Lymphoma; HL: Hodgkin Lymphoma; SCM: Stem-cell mobilization; BM: Bone marrow; PLX: Plerixafor. * Either Hb<13 g/dl or Plt<170.000/mmc and WBC<5000/mmc. ** Concurrent use of upfront Plx and chemo-based SCT was not frequent in our sample and thus the reported probability may be inaccurate